

# Middlesex University Research Repository

An open access repository of

Middlesex University research

<http://eprints.mdx.ac.uk>

Baluch, Bahman ORCID logoORCID: <https://orcid.org/0000-0003-0650-4421> (2005) Persian orthography and its relation to literacy. In: Handbook of Orthography and literacy. Joshi, R. Malatesha and Aaron, P. G., eds. Lawrence Erlbaum Associates, London. ISBN 0805846522.  
[Book Section]

This version is available at: <https://eprints.mdx.ac.uk/7649/>

## Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author's name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:

[eprints@mdx.ac.uk](mailto:eprints@mdx.ac.uk)

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: <http://eprints.mdx.ac.uk/policies.html#copy>

# 22

## Persian Orthography and Its Relation to Literacy

Bahman Baluch  
*Middlesex University*

Persian orthography, which is a modified version of the Arabic script, is used for transcribing the Persian (Farsi) language, which is the major language spoken in Iran. Persian is also one of the two (Farsi and Urdu) major languages spoken in Afghanistan, and the main language in Tajikistan, a former central Asian republic of the former Soviet Union. However, the Persian spoken in these countries and the script used to transcribe the spoken language, particularly in Tajikistan, have been influenced by local factors and borrowed words. The focus of this chapter is on the Persian spoken in the present-day Iran and its relationship with the orthography, henceforth referred to as Persian orthography. In particular, the emphasis is on how literacy acquisition by Persian beginner (and skilled) readers may be affected by peculiarities of Persian orthography. Arguably, very little systematic research has been conducted on cognitive processes involved in the reading of Persian. It is hoped that this article will stimulate such research. After presenting factors that influence literacy acquisition in Persian, we take up the question of whether there should be changes to Persian orthography.

### INTRODUCTION

This chapter is divided into two sections: The first section deals with the Persian orthography and the nature of its phonological and morphological structure. These features are reviewed from a historical perspective followed by a discussion of how these historical changes have affected modern Persian. The second section deals with literacy acquisition and the nature of reading problems encountered by beginner and skilled readers. More specifically, we also examine the relationship between the written and spoken forms of language, potential difficulties the Persian orthography might present for the reader in the process of literacy acquisition, and the development of skilled reading. In the concluding part of this section, the question of whether introducing changes in the Persian orthography would minimize some of the difficulties being experienced by readers is raised.

## PERSIAN, ITS ORTHOGRAPHY, MORPHOLOGY, AND PHONOLOGY

Modern Persian is derived from Indo-Iranian, one of the branches of the Indo-European languages. As shown in Fig. 22.1, the Indo-European language group split into Indo-Iranian and Indic language groups, from which most of the languages of India are derived. This development is estimated to have taken place around 1500 B.C.E. (Khanlari, 1979, 1995).

The old Persian, dating back to around 551 B.C.E. was written in cuneiform, the wedge-shaped characters used throughout much of the Middle Eastern countries. Around the historical period known as the middle Persian (around 331 B.C.E.) the Persians created their own writing system, known as *Pahlavi*, which remained in use until the Islamic conquest of the seventh century. What is referred to as modern Persian script is a transcription of Persian by a modified version of the Arabic script and therefore dates from the seventh century, marked by the Arabic conquest of Persia. Modern Persian is spoken by over 50 million people in Iran and another 5 million in Afghanistan. In Iran it is generally referred to as Farsi, but in Afghanistan as Dari. A variety of Persian called Tajik is spoken in Tajikistan; however, it is written in the Cyrillic alphabet (Khanlari, 1979, 1995).

### The Persian Alphabet and Its Peculiarities

The Persian alphabet comprises 32 letters, the original 28 Arabic letters and an addition of four letters that represent Persian phonemes that are not represented by Arabic letters. The additional four letters are graphically identical to Arabic letters but differ only with the addition of dots to the Arabic letters. Examples are the Arabic letters *ر*/re/ and *ز*/zeh/ and the invented Persian letter *چ*/jeh/. Furthermore, the addition of a stroke on the top of the Arabic letter *ك*/keh/ and the Persian invented letter *گ*/geh/ distinguishes these letters from the original Arabic letters. However, the rules of transcription and letter shapes for the invented letters follow that of Arabic writing. Indeed, it is a noticeable feature of the Arabic alphabet and invented Persian letters that there are only eight basic forms for letters. Other letters are simply a variant of one of these basic forms, with the difference of dot(s) in almost all cases. The following is the list of

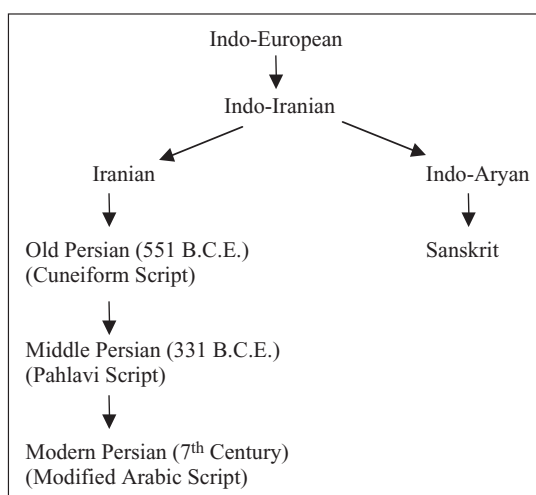


FIG. 22.1. Schematic presentation of the development of historical roots of the Persian language and script.

basic letter forms on which other letters of the alphabet are created with the addition of dot(s) above or below the basic letter shape: ب /beh/, ح /he/, د /dâl/, ز /re/, س /se/, ص /sât/, ع /eyin/, ط /tâ/.

Persian, like all Semitic scripts (Arabic and Hebrew), is written from right to left. Most letters are written in a joined fashion, rather like English cursive handwriting. However, some letters, depending on their position in the word, are never joined to a following letter. Although there are no uppercases or lowercases in Persian letters, there are many letters that have different shapes depending on their position in the word and whether or not they should be written connected to another letter. For example, the letter غ /geh/ takes on different forms depending on its position as initial غ or middle غ or final غ. Word-final forms mark the word boundaries. Even though the transcription of Persian letters in cursive format is helpful in marking word boundaries, readers may experience difficulty in deciding word boundaries when they encounter letters that do not join other letters. The letters ا /alef/, د /dâl/, ذ /zâl/, ر /re/, ز /ze/, ژ /jeh/, and و /vâv/ have only one form, regardless of their position within the word. Like English, most words are separated by a space. However, in view of the cursive form of written Persian, if the first word ends in one of the characters that is left unjoined to a preceding letter (e.g., the word /mozd/ [wages], transcribed as مزد, the end of the word may not be as predictable as most other words in the text.

## The Vowels and Diacritics

The number of vowels has been reduced from eight in old Persian to six (â, a, e, i, o, u) in modern Persian. Three vowels (a, e, and o), generally known as short vowels, are simply represented by diacritic superscript or subscript marks attached to the letters of the alphabet. The other three vowels are the long vowels (â, î, û) and are conveyed by letters of the alphabet (see Fig. 22.2 for examples of Persian words with vowel letters and diacritics).

There is a direct, one-to-one relationship between letters of the alphabet and phonemes in Persian. Moreover, Persian script, insofar as grapheme–phoneme correspondences are concerned, in its fully vowelised format is a highly regular orthography. For this reason, it is similar to Roman orthographies such as Turkish. However, similar to English, in Persian a phoneme may be presented by more than one letter of the alphabet. In addition, in practice, diacritic vowels are used in writings used by beginning readers and religious writings; vowels are almost always omitted from general text. Long vowels are never omitted from written text (see Fig. 22.2 for examples). This creates ambiguities, namely phonological, semantic, and visual–orthographic (Baluch, 1990, 1992).

Phonological and semantic ambiguity results when the reader is faced with a string of consonantal letters. For example, the consonantal string /krm/ can be pronounced with different vowel combinations resulting in five possible pronunciations and meanings /kerm/ [worm], /karam/ [generosity], /kerem/ [cream], /krom/ [chrome], and /karm/ [vine]. The manner in which a reader may eventually retrieve the correct pronunciation (and meaning) has been the subject of the research that is discussed in a subsequent section of this chapter. Adding to the confusion is the

	Persian Spelling	Transcription	Meaning
Vowel letter/â/	سال	/sal/	year
Diacritic /a/	سزد	/sard/	cold
No diacritic	سرد	/srd/	cold

FIG. 22.2. Examples of Persian words with vowel letters, diacritic vowels, and with no diacritics specified.

many meanings that one could still infer by adding the (absent) stress assignment (see next subsection). For example, in the preceding string, depending on stress assignment, one may read /karam/ to mean ‘generosity’ or with the stress assignment on the /a/ to mean ‘I am deaf’.

Another possible source of ambiguity is visual. The absence of vowels in written Persian makes the words somewhat shorter. Thus there are fewer visual cues as to the identity of the word. Indeed, research on English has shown a direct relationship between word length and visual word recognition (see Weekes, 1997).

### Stress Assignment and Colloquial Ambiguity

The rules of stress assignment and intonations in Persian have been discussed elsewhere (see, e.g., Lazard, 1992). What is important to note is that, although understanding these rules may contribute to phonological disambiguation of written Persian, by itself it falls short when faced with problems caused by colloquial ambiguity. The change in the spoken version, what is known as a ‘Tehrani accent’, has resulted in complete changes in stress assignments to a large corpus of words. For example, the word /miguyand/ [they say] undergoes a complete vowel change in colloquial language by being pronounced as /migan/; the stress assignment to the verb /xastan/ [to want] when used in written correct format would be /mixaham/ [I want]; with a Tehrani accent the stress is on the initial and final syllables resulting in /mixam/. Labelled as diglossia, the possible psychological impact of this dissociation between written and spoken language has been the subject of investigation (see, e.g., Hudson, 1992).

### The Persian Morphological System

In spite of the fact that a considerable portion of the Persian lexicon is derived from Arabic roots, including the Arabic plural patterns, Persian morphology and orthography do not match perfectly. Persian morphology is an affixal system consisting mainly of suffixes and a few prefixes. Thus when it relates to Arabic words, two kinds of spellings and morphological processes are encountered by the reader. Figure 22.3 shows how the Arabic root system (a consonant string) is used for deriving nouns by the insertion of certain vowel patterns and the way it differs from a similar process in Persian.

In the preceding example, the Arabic plural form for ketâb is /kotob/, obtained by the root derivation system. In Persian, the plural for the lexical word /ketâb/ can be given as it is in (/kotob/), or it can be obtained just by the addition of the Persian plural morpheme /hâ/ = /ketâb+hâ/. Any new Persian word, however, can be pluralized only by the addition of the plural morpheme. In addition, because the plurals formed by the Arabic morphological system constitute a small portion of the Persian vocabulary (about 5%; Khanlari, 1995), it is not necessary to include them in the morphology; they are listed instead in the dictionary as irregular forms.

Root letter string	Arabic word/ meaning/spelling	Arabic plural/spelling
/k-t-b/	/ketâ b/ “book” کتاب	كتب
Root letter string	Persian word/meaning/spelling	Persian plural/spelling
/k-t-b/	/ketâ b/ “book” کتاب	کتابها

FIG. 22.3. Example of the root consonant string /ktb/ and the morphological processes in Arabic and Persian.

Persian word/ meaning	with suffix y/meaning	
/mard/ “man”	/mardy/	“a man” Noun+Indefinite marker
	/mardy/	“the man” Noun+Relative clause linking enclitic
	/mardy/	“you are man” Noun+Copula (second, singular)
	/mardy/	“manliness” Noun+derivational morpheme

FIG. 22.4. Examples of morphological ambiguity in Persian depending on stress assignment.

Although the presence of suffixes may seem more helpful to Persian readers in the absence of short vowels, there are still many examples of morphological ambiguity. For example, the root word written as */mrd/*, when pronounced with the */a/*, reads as */mard/* [man]. When the suffix *y* is added and read with the short vowel, it reads as */mardy/* [a man] (see Fig. 22.4 for examples). In addition, the same spelling also can give rise to the morphological ambiguities depending on stress assignment, as shown with an example in Fig. 22.4.

### The Role of the *Ezafeh* Morpheme

An issue not discussed in the past by researchers on Persian is the difficulties that readers may encounter in determining phrase boundaries, especially the boundaries of noun phrases. Determining phrase boundaries is difficult because Persian is a verb-final language and there are no markers to distinguish the subject or the objects in a sentence (with the exception of the specific object marker *râ*). For example, in English one adds the suffix to the phrase ‘John’s book’ to indicate that the book belongs to John. In Persian, a phrase to indicate that a book belongs to, say, Ali, would be written as */ktâb/ /ali/* and reads as ‘book ali’. What is missing is the *ezafeh* morpheme, usually an unwritten vowel pronounced in the latter example as */e/*, which is a short vowel and is therefore not written in text. The *ezafeh* only appears in written text after the vowels */â/* and */u/*. Thus, without any clear markers to determine the phrase boundaries, and without the *ezafeh* (in most cases) to link the phrase constituents, Persian readers may face difficulty in determining meaning and end of phrases.

### Summary

In today’s Iran, a modified version of Arabic script is used to transcribe the spoken language. Both Arabic and modified Arabic used to transcribe Persian are argued to be phonologically opaque because short vowels are omitted from the script. It may also be argued to be visually opaque as most letters differ only with the presence or absence of dots. However, although in then fully vowelised format Persian and Arabic are highly regular in the direction of grapheme–phoneme correspondences, the same is not true for phoneme–grapheme relationships. Thus there is considerable ambiguity that is due to which letter should be used in relation to a specific phoneme.

Adding to these there are further problems that are due to the importation of Arabic script to fit an Indo-European language. Most noticeable are the irregularities of the imported Arabic words and the use of both Persian and Arabic morphological processes in the derivation of various variants of the root morpheme. Moreover, in view of the grammatical nature of Persian and the absence of morpheme *ezafeh* there is clear difficulty in determining phrase boundaries. Added to this is also the issue of the possible impact on literacy dissociation of colloquial Persian, the ‘Tehrani accent’, and the standard Persian.

## THE IMPACT OF PERSIAN ORTHOGRAPHY ON LITERACY ACQUISITION

### Introduction

In this section the issue of cognitive processes involved in literacy acquisition of Persian is reviewed. The question pursued here is the extent to which the peculiarities outlined in the previous section, in relation to Persian orthography, affect literacy acquisition. According to the recent statistics by the office of National Statistics of the Islamic Republic of Iran, the literacy rate is 76% and the illiterate population stands at 10.6 million.

Of the main orthographic and phonological factors that may be thought to affect literacy in Persian one may identify the grapheme–phoneme regularity, the phoneme–grapheme ambiguity, and the absence of short vowels in written text. These peculiarities of Persian have been the subject of limited investigation (Arab-Moghaddam & Senechal, 2001; Baluch, 1990, 1993; Gholamain & Geva, 1999). Other factors such as the preponderance of dots to differentiate letters, letter similarity, and variety of letter shapes, the omission of stress symbols, the absence of the marker *ezafeh* as indication of phrase boundary, the morphological structure of Persian, and the issue of diglossia have yet to be investigated.

### Grapheme–Phoneme Regularity and Literacy

One of the most noticeable aspects of Semitic scripts, and in particular Persian orthography, is the regularity of grapheme- (letter-) to-sound relationships. The issue of whether such regularity may have an impact on learning to read has been investigated in greater length among orthographies other than Persian that differ in the manner in which they represent the phonology of the spoken language (see, for example, Geva, 1995). The term ‘orthographic depth’ is used to distinguish orthographies on a continuum ranging from shallow to opaque (Baluch, 1993; Baluch & Besner, 1991; Frost, Katz, & Bentin, 1987). Implicit in this is the assumption that some alphabetic scripts labelled as shallow, like Persian, Hebrew and Arabic, and the Roman scripts like Turkish and Italian, have simple grapheme–phoneme relationships whereas other scripts like English, labelled as deep, have more complicated grapheme–phoneme structures. In terms of literacy acquisition, the argument goes that shallow orthographies are easier to acquire than deep orthographies because ‘the development of word-based processes in different languages might vary as a function of orthographic regularity’ (Gholamain & Geva, 1999, p. 184). This script-dependent hypothesis has found support from various studies on deep and shallow orthographies that have shown that shallow orthographies have lower incidents of reading disability than do deep orthographies (e.g., Goswami, Schneider & Scheurich, 1999; Lindgren, de-Renzi, & Richman, 1985; Wydell & Butterworth, 1999). There is also evidence that word recognition in terms of oral naming is easier in shallow orthographies such as vowelised Hebrew (Geva, 1995), and Turkish (Oney & Golden, 1984; Oney, Peter, & Katz, 1997; Raman, 1999).

However, an alternative hypothesis, the ‘central-processing hypothesis’, maintains that skilled reading depends on the efficient functioning of working memory, naming speed, and lexical processes. If these cognitive and linguistic skills are deficient, the individual is likely to experience difficulties in the acquisition of literacy skills, regardless of the orthography involved (e.g., Bowers, 1995; Bowers, Golden, Kennedy, & Young, 1994). There is of course a possibility of a compromise between these two positions, namely, that there are some basic mechanisms that have to be in place for efficient functioning of reading performance, such as an efficient working memory. What enhances the functioning of this system is how transparent

the orthography of a given language is (see, for example, Baluch, 1990; Geva, 1995; Baluch & Danaye-Tousie, 2005).

### Evidence From Persian

An extensive study by Gholamain and Geva (1999) examined the linguistic, cognitive, and basic reading skills of 70 Persian children from immigrant families in Canada in Grades 1 to 5. These children were learning to read concurrently English and Persian. Gholamain and Geva (1999) argued that, because Persian has a reliable one-to-one grapheme–phoneme correspondence, it should be much easier for the children to master the grapheme–phoneme correspondence rules and therefore be able to read and decode Persian words with a relatively greater degree of accuracy in the early grades than might be the case for reading English. If confirmed, this hypothesis, of course, would support the script-dependent hypothesis. Indeed, in line with this prediction, it was found that with only 3 hr of Persian literacy instructions per week, once the children had mastered the rules they were able to decode even unfamiliar complex Persian words. However, Gholamain and Geva (1999) also reported data in support of the central-processing hypothesis in that they found that children who performed better on measures of reading and cognitive skills in English, their primary language, were more likely to perform better in Persian, their second language. In particular, the role played by working memory and rapid automatised naming was highly correlated between the two languages.

There is, however, a methodological question regarding the research just described and indeed any research using bilingual participants, namely, there are the possible effects of learning strategies used in one language being generalised to a second language. The immigrant Persian children studied in Gholamain and Geva (1999) were learning Persian as a second language in Canada. For example, Mumtaz and Humphreys (2001) reported that Pakistani children learning Urdu (which, like Persian, is transcribed by a modified version of the Arabic script) and English make errors that are more indicative of over-generalisation of strategies used in one language to a second language.

There is very little, if any, published work that has examined the development of reading skills among monolingual Persian children learning to read and write in their homeland. Two unpublished master's dissertations are the only sources of data that could shed some light on reading performance of monolingual Persian children. Amini (1997) studied 120 normal and dyslexic first- and second-grade elementary school children on a battery of 20 cognitive tests. Amini concluded that there are greater similarities between factors affecting poor reading in Persian than those reported in the literature on reading English. She argued that her data on Persian children are more in line with the central-processing hypothesis. The problem with Amini's study is that there are no reports on how the batteries of tests used in her study were developed and validated for her investigation.

Shirazi's (1996) master's dissertation on monolingual Persian children may also be taken as a further support for the central-processing hypothesis. Shirazi's research was aimed at 67 Persian children, 35 girls and 32 boys 6–7 years old.

Shirazi administered the tests of rapid naming, phonological awareness, verbal working memory span, and oral reading to first-grade children in a school in Tehran during the first month of their reading instruction, and later tested the same children 3 months later. Shirazi reported that there is a significant relationship between phonological awareness and oral reading speed. She further argued that accuracy in reading and improvement in phonological awareness are highly related to each other. Moreover, those children performing significantly better on phonological awareness were also ranked higher on their reading performance by their teacher. Although Shirazi's sample is small and probably not representative of the children in Iran, it



is nevertheless a good attempt to develop a Persian version of cognitive tests used in western countries to conduct research on Persian readers.

### Phoneme–Grapheme Ambiguity and Literacy

The regular grapheme–phoneme correspondences in Persian may facilitate oral reading performance; however, the same is not true about the ability to spell. This is because, in Persian, the same phoneme may be represented by more than one grapheme (polygraphy). Thus the expectation is that, in order to decide on the use of the correct grapheme for the word’s spelling, Persian readers may have to rely on their lexical knowledge. Therefore a major problem encountered in children’s spelling is letter substitution.

Azzam (1989) reported many errors by Arab children in using the correct letters when engaged in a spelling task; for example, the errors made on the choice between the letters *z* and *ze*, both of which are pronounced /ze/. One reason for the errors made by Arab children may be due to the extensive reliance on phonological strategies when dealing with the spelling task. This is because at the grapheme–phoneme level there is a very transparent relationship in Arabic. Hence children experiencing regularities at grapheme–phoneme level for oral reading may be less likely to use their lexical knowledge for spelling words and hence the kind of errors reported in Azzam’s (1989) study. The question is whether the same is true for readers of Persian. Cossu (1995) examined possible discrepancies between reading and spelling strategies for Italian children. Cossu (1995) argued that, because Italian is transparent in both grapheme–phoneme and phoneme–grapheme correspondences, this should encourage parallel strategies and similar level of performance. Cossu (1995), however, reported that reading accuracy is significantly better than spelling accuracy for first- and second-grade Italian children. This indicates that although Italian is a very transparent script it does not necessarily encourage parallel strategies in reading and spelling.

Oney et al. (1997) compared reading and spelling performance of English and Turkish readers in first, second, and fifth grades and found that orthographic transparency determines the degree to which readers use phonology during word recognition and suggests that readers become less dependent on phonological mediation with experience and that this reduction is more rapid for readers of opaque orthographies. Thus for reading English the fact that both grapheme–phoneme and phoneme–grapheme correspondences are opaque there is a greater reliance on non-phonological strategies, whereas for readers of transparent Turkish this non-reliance may take longer.

### Evidence From Persian

In an extensive study, Arab-Moghaddam and Senechal (2001) examined orthographic and phonological processing skills in bilingual Persian–English children, targeting a relatively large sample of 55 bilingual children living in Canada. They argued that because Persian is polygraphy but not polyphony, whereas English is both polygraphy and polyphony, it should encourage different phonological and orthographic strategies in literacy acquisition of the two languages. In the case of oral reading, Arab-Moghaddam and Senechal (2001) found that both phonological and orthographic skills were predictors of good performance in both English and Persian. However, Arab-Moghaddam and Senechal (2001) reported that, first, Persian children were better able to spell words in English than they did in Persian, and, second, the orthographic skills were a key predictor of their spelling ability. The argument put forward was that the nature of the Persian orthography encourages children to adopt different strategies when spelling words. Spelling words in Persian is inefficient if an analytic strategy is used, and

perhaps the realisation that this is an inappropriate strategy comes in at a later stage compared with that of English.

### Phonological Ambiguity Due to the Omission of Short Vowels

As outlined in the section titled “The Vowels and Diacritics,” the short vowels in Persian are used only for beginner readers. Thus the reader is faced with a string of mainly consonantal spelling (Baluch, 1990, 1992). The expectation is that, by the time the diacritic marks are omitted, the reader has developed appropriate strategies for cognitive processes in reading and a well-established visual orthographic lexicon (Baluch, 1992). However, the fact that vowels are absent from written text may indeed cause ambiguity of various degrees for readers of Persian.

### Evidence From Persian

Baluch and Shahidi (1991) studied naming of words with consonantal spelling and matched words with vowel letters by Persian children with the mean age of 8.4 years. The children were taught to read under the traditional system of first learning to read words with the use of diacritics followed by omission of diacritics. The results showed that there were significantly more errors made to opaque words (e.g., بچه /bch/, pronounced as /baceh/ [child]) compared with transparent words (e.g., بازی /bâzi/ [play]). Moreover, the time taken to name a list of words with consonantal spelling was significantly slower than the time in naming a list of words with vowel letter spellings.

Baluch (1990, 1993) reported similar findings when the oral naming of consonantal words and vowel letter words was examined by skilled adult Persian readers (see Table 22.1).

For vowel letter spellings, even of a high-frequency nature, there was a 65-ms difference in oral naming than for consonantal spellings, a difference that was even greater for low-frequency words. Moreover, a consonantal word with multiple meanings was named significantly slower than a consonantal word with a unique spelling. These findings demonstrate significant difficulty in naming consonantal spelling by use of phonological processes.

**TABLE 22.1**  
Mean Reaction-Time Latencies, Standard Deviations (in Parentheses), and Error Percentages to High- and Low-Frequency Persian Words and to Opaque-Ambiguous and Opaque-Unique Persian Words

<i>Spelling Type</i>	<i>Opaque</i>	<i>Transparent</i>	<i>Difference</i>
<b>High Frequency</b>			
Mean ( <i>SD</i> )	615 (137)	550 (122)	65
Error (%)	5.8	2.5	3.3
<b>Low Frequency</b>			
Mean ( <i>SD</i> )	652.5 (151)	565.9 (110)	86.6
Error (%)	14.6	4.16	10.4
	<i>Opaque (Polysemous)</i>	<i>Opaque (Monosemous)</i>	<i>Difference</i>
Mean ( <i>SD</i> )	600 (124)	551 (94)	49
Error (%)	4.16	2.8	1.36

Abu-Rabia (1997) came to similar conclusions in relation to reading Arabic. Abu-Rabia (1997) investigated the effect of vowels on reading accuracy of poor and skilled native Arabic readers in Israel. The materials used were narrative stories and newspaper articles. His subjects read Arabic narrative stories and newspaper articles under four reading conditions: vowelised text, unvowelised text, vowelised word naming, and unvowelised word naming. The results showed that vowels and text contexts were important variables that facilitate word recognition in poor as well as skilled readers in Arabic orthography. Vowels speeded word naming and context facilitated disambiguation of polysemous words.

The question may then be raised as to the costs and benefits of having the diacritics omitted from text for both beginning and skilled readers (Baluch, 1992).

### Other Possible Factors That Affect Literacy Acquisition of Persian

Greater letter similarity and confusion in distinguishing between different letter shapes and their identity, the absence of *ezafeh* as markers of word boundary, and diglossia in Persian, that is, the strong dissociation of standard written Persian and colloquial Persian (Tehrani accent), are other sources of ambiguity encountered by Persian readers.

The greater letter similarities of Arabic have been noted by Azzam (1989). She reported that Arab readers have difficulty distinguishing between different letters (e.g., ز and ط) and between consonantal and vowel letters (vowel diacritic و and vowel letter و). Moreover, Azzam reported on the errors made by Arab children in reading words with letters that change shape depending on their position in the word. She reported that children were confused when reading long vowel /a/ with alef ا or yeh ی for instance, children may read متا rather than متى.

In relation to diglossia, Abu-Rabia (2000) examined the influence of exposure to literary Arabic on reading comprehension in Arabic-speaking children. He concluded that exposure of preschool children to literary Arabic in diglossic situations enhances their reading ability in the first and second grades. Ravid (1996) came to similar conclusions when examining reading performance of both children and adult readers of Hebrew. There is no reported research of a similar kind in Persian.

## CONCLUDING REMARKS

In short, the arguments put forward in this section of this chapter suggest that perhaps the only beneficial aspect of Persian orthography, insofar as literacy acquisition is concerned, is a relatively easier task enjoyed by Persian beginning readers in oral naming. The ambiguity at phoneme–grapheme level and the absence of short vowels, in addition to a host of yet uninvestigated factors associated with Persian orthography, may be other significant sources that affect literacy acquisition.

A final note that may be made here is whether some changes should be introduced into Persian orthography. More than a couple of decades ago Nickjoo (1979) highlighted some of the key features of the Persian scripts, namely, letter similarities, phoneme–grapheme ambiguity, absence of vowels, and many variants of the position-dependent letters. Nickjoo (1979) argued that such peculiarities of written Persian have implications for literacy and argued for the abolition of the Persian alphabet and the creation of a Latinised version of Persian. Nevertheless, Nickjoo (1979) acknowledged that the political and religious considerations may be a serious hindrance for any such reform. Similar political and religious considerations were also heavily present in relation to the transformation of Turkish writing system. However, in 1931, Kamal Atatürk, the father of modern Turkey, with the help of the slogan “the

Turkish language has been a prisoner for centuries and is now casting off its chains” set in motion a change of Turkish script from Arabic to Roman (Raman, 1999). Moreover, an attempt was made to make the correspondence as reliable and as transparent as possible. Although the political and religious debates may be the key factors affecting decisions on whether to change the Persian alphabet, there is also an interesting possibility for a more scientific argument to enter the equation. Because a generation of Turkish readers have now experienced reading in the Romanised script, a comparison of literacy acquisition by both beginning and skilled readers of Turkish and Persian language may shed more light as to the possible impact of orthography change on literacy acquisition.

## REFERENCES

- Abu-Rabia, S. (1997). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers. *Reading and Writing*, 9, 65–78.
- Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing*, 13, 147–157.
- Amini, M. (1997). *Differences between normal and dyslexic children in reading and writing Persian texts*. Unpublished master’s dissertation, College of Rehabilitation Sciences, Tehran, Iran.
- Arab-Moghaddam, N., & Senechal, M. (2001). Orthographic and phonological processing skills in reading and spelling in Persian/English bilinguals. *International Journal of Behavioral Development*, 25, 140–147.
- Azzam, R. (1989). Orthography and reading of the Arabic language. In P. G. Aaron & R. M. Joshi (Eds.), *Reading and writing disorders in different orthographic systems* (pp. 203–218). Dordrecht, The Netherlands: Kluwer Academic.
- Baluch, B. (1990). Word recognition in Persian. *Baranoosh Scientific Quarterly*, 1(1), 3–6.
- Baluch, B. (1992). Reading with and without vowels: What are the psychological consequences? *Journal of Social and Evolutionary Systems*, 15, 95–104.
- Baluch, B. (1993). Lexical decisions in Persian: A test of the orthographic depth hypothesis. *International Journal of Psychology*, 28, 19–27.
- Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology, Learning, Memory and Cognition*, 17, 644–651.
- Baluch, B., & Danaye-Tousie, M. (2005, January). *Developmental dyslexia as a function of opaqueness and transparency of a word’s spelling: Evidence from Persian*. Presented at the meeting of the Experimental Psychology Society, London.
- Baluch, B., & Shahidi, S. (1991). Visual word recognition in beginning readers of Persian. *Perceptual and Motor Skills*, 72, 1327–1331.
- Bowers, P. G. (1995). Tracing symbol naming speed’s unique contributions to reading disabilities over time. *Reading and Writing*, 7, 189–216.
- Bowers, P. G., Golden, J., Kennedy, A., & Young, A. (1994). Limits upon orthographic knowledge due to processes indexed by naming speed. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge, 1: Theoretical and developmental issues* (pp. 173–218). Dordrecht, The Netherlands: Kluwer Academic.
- Cossu, G. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non parallel processes? *Reading and Writing*, 7, 9–22.
- Frost R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 104–115.
- Geva, E. (1995). Orthographic and cognitive processing in learning to read English and Hebrew. In I. Taylor & D. R. Olson (Eds.), *Scripts and literacy: Reading and learning to read alphabets, syllabaries and characters* (pp. 277–291). Dordrecht, The Netherlands: Kluwer Academic.
- Gholamain, M., & Geva, E. (1999). Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning*, 49, 183–217.
- Goswami, U., Schneider, W., & Scheurich, B. (1999). Picture naming deficits in developmental dyslexia in German. *Developmental Science*, 2, 53–58.
- Hudson, A. (1992). Diglossia: A bibliographic review. *Language in Society*, 21, 611–674.
- Khanlari, P. N. (1979). *The history of Persian language (Vol. 1)*. New Delhi, India: New Delhi Press.
- Khanlari, P. (1995). *Tarikh-e Zaban-e Farsi* [History of the Persian Language]. Tehran, Iran: Simorgh Press.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 2, 293–323.
- Lazard, G. (1992). *A Grammar of Contemporary Persian*. Costa Mesa, CA: Mazda Publishers.

- Lindgren, S. D., de-Renzi, E., & Richman, L. C. (1985). Cross-national comparisons of developmental dyslexia in Italy and the United States. *Child Development*, 56, 1404–1417.
- Mumtaz, S., & Humphreys, G. W. (2001). The effects of bilingualism on learning to read English: Evidence from the contrast between Urdu–English bilingual and English monolingual children. *Journal of Research in Reading*, 24, 113–134.
- Nickjoo, M. (1979, May). A century of struggle for the reform of the Persian script. *The Reading Teacher*, 926–929.
- Oney, B., & Goldman, S. R. (1984). Decoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme-phoneme correspondences. *Journal of Educational Psychology*, 76, 557–567.
- Oney, B., Peter, M., & Katz, L. (1997). Phonological processing in printed word recognition: Effects of age and writing system. *Scientific Studies of Reading*, 1, 65–83.
- Raman, I. (1999). *Single word naming in a transparent orthography*. Unpublished Ph.D. thesis, Middlesex University, England.
- Raman, I., Baluch, B., & Sneddon, P. (1996). What is the cognitive systems preferred route for deriving phonology from print? *European Psychologist*, 1, 221–227.
- Ravid, D. (1996). Accessing the mental lexicon: Evidence from incompatibility between representation of spoken and written morphology. *Linguistics*, 34, 1219–1246.
- Shirazi, T. S. (1996). Phonological awareness as an important predictor of reading acquisition in Persian speaking children. Unpublished master’s dissertation, University of Rehabilitation and Social Welfare, Tehran, Iran.
- Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A, 439–456.
- Wydell, T. N., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, 70, 273–305.